AbstractID: 12981 Title: On the Correlation of Conventional IMRT QA Metrics to Clinical Impact in Per-Patient Dose QA

Purpose: Per-patient IMRT QA is often performed by comparing measured planar phantom dose to calculated planar phantom dose, per beam. The most common analysis metric is percent passing 3%/3mm (DTA). Similarly, TG-119 reports utilizes a metric of 3%/3mm (Gamma) when discussing IMRT commissioning. In this work, we analyze the correlation of conventional per-beam IMRT QA metrics to: 1) patient dose comparisons (3D %diff/DTA), and 2) important clinical dose endpoints in volumetric regions-of-interest.

Method and Materials: For 24 Head/Neck IMRT plans, an accurate TPS beam model was error-induced (four unique ways) to simulate IMRT calculation and/or delivery errors. The error-free beam models were used to generate “virtual planar QA measurements” and analyzed in conventional ways against the error-induced beams. The error-free 3D patient doses were compared against the error-induced 3D patient doses. This method produced a robust simulation (high precision, high data density of 3D patient dose) which is necessary given the practical inability to measure dose inside complicated patient anatomy. The 96 error-induced plans were compared to the error-free plans to quantify the correlation of conventional Planar QA metrics with relevant clinical endpoints such as: max cord dose, mean parotid dose, and D95 of the principal CTV.

Results: No correlation was found between 3%/3mm conventional IMRT QA criteria and all clinical endpoints studied (R-squared values of: 0.0314, 0.0272, 0.0899, and 0.0248 for predicting errors in the cord max, left parotid mean, right parotid mean, and CTV D95). There was also no correlation when considering 2%/2mm and 1%/1mm metrics. In addition, there were many examples of both false positives and false negatives using conventional IMRT QA methods/metrics to predict accurate patient dose.

Conclusion: The location and magnitude of the errors was far more important than the quantity of the errors, and moving to patient ROI-based error simulations is strongly recommended.